

High-Capacity Quantum Associative Memories

M. Cristina Diamantini*

*INFN and Dipartimento di Fisica, University of Perugia,
via A. Pascoli, I-06100 Perugia, Italy*

Carlo A. Trugenberger[†]

SwissScientific, chemin Diodati 10, CH-1223 Cologny, Switzerland

(Dated: June 4, 2015)

Abstract

We review our models of quantum associative memories that represent the “quantization” of fully coupled neural networks like the Hopfield model. The idea is to replace the classical irreversible attractor dynamics driven by an Ising model with pattern-dependent weights by the reversible rotation of an input quantum state onto an output quantum state consisting of a linear superposition with probability amplitudes peaked on the stored pattern closest to the input in Hamming distance, resulting in a high probability of measuring a memory pattern very similar to the input. The unitary operator implementing this transformation can be formulated as a sequence of one-qubit and two-qubit elementary quantum gates and is thus the exponential of an ordered quantum Ising model with sequential operations and with pattern-dependent interactions, exactly as in the classical case. Probabilistic quantum memories, that make use of postselection of the measurement result of control qubits, overcome the famed linear storage limitation of their classical counterparts because they permit to completely *eliminate crosstalk and spurious memories*. The number of control qubits plays the role of an inverse fictitious temperature, the accuracy of pattern retrieval can be tuned by lowering the fictitious temperature under a critical value for quantum content association while the complexity of the retrieval algorithm remains polynomial for any number of patterns polynomial in the number of qubits. These models solve thus the capacity shortage problem of classical associative memories, providing a *polynomial improvement* in capacity. The price to pay is the probabilistic nature of information retrieval.

PACS numbers: 03.67.-a, 03.67.Lx, 64.70.Tg

*Electronic address: cristina.diamantini@pg.infn.it

†Electronic address: ca.trugenberger@bluewin.ch

I. INTRODUCTION

There is a growing consensus that the fundamental mechanism of human intelligence is simply pattern recognition, the retrieval of information based on content association, albeit repeated in ever increasing hierarchical structures [1]. Correspondingly, pattern recognition in machine intelligence [2] has made enormous progress in the last decade or so and such systems are now to be found in applications ranging from medical diagnosis to facial and voice recognition in security and digital personal assistants, the latest addition to the family being self-driving cars. On the other side, the last two decades have seen the birth of, and an explosion of research in a new information-theoretic field: quantum information theory and quantum computation [4]. This chapter deals with quantum pattern recognition, with particular emphasis on models that are both accessible to detailed analytical treatment and efficiently implementable within the framework of the quantum circuit model.

Pattern recognizers, which go also under the name of *associative memories* (or more precisely autoassociative memories), are fundamentally different than von Neumann or Turing machines [3], which have grown into the ubiquitous computers that permeate our information society. Computation is not sequential but, rather, based on collective phenomena due to interactions among a large number of, typically redundant, elementary components. Information is not address-oriented, i.e. stored in look-up tables (random access memories, RAMs) but, rather, distributed in often very complex ways over the connections and interactions parameters. In traditional computers information is identified by a label and stored in a database indexed by these labels. Retrieval requires the exact knowledge of the relevant label, without which information is simply not accessible. This is definitely not how our own brain works. When trying to recognize a person from a blurred photo it is totally useless to know that it is the 16878th person you met in your life. Rather, the recognition process is based on our strong power of association with stored memories that resemble the given picture. Association is what we use every time we solve a crossword puzzle and is distinctive of the human brain.

The best known examples of pattern recognizers are neural networks [5] and hidden Markov models [6], the Hopfield model [7] (and its generalization to a bidirectional associative memory [8]) being the paradigm, since it can be studied analytically in detail by the techniques of statistical mechanics [5, 9]. The great advantage of these architectures is that

they eliminate the extreme rigidity of RAM memories, which require a precise knowledge of the memory address and, thus, do not permit the retrieval of incomplete or corrupted inputs. In associative memories, on the contrary, recall of information is possible also on the basis of partial knowledge of its content, without knowing a precise storage location, which typically does not even exist. This is why they are also called “content-addressable memories”.

Unfortunately, classical associative memories suffer from a severe capacity shortage. When storing multiple patterns, these interfere with each other, a phenomenon that goes under the name of crosstalk. Above a critical number of patterns, crosstalk becomes so strong that a phase transition to a completely disordered spin glass phase [10] takes place. In this phase there is no relation whatsoever between the information encoded in the memory and the original patterns. For the Hopfield model, the critical threshold on the number p of patterns that can be stored in a network of n binary neurons is $p_{max} \simeq 0.138 n$ [5]. While various possible improvements can be envisaged, the maximum number of patterns remains linear in the number of neurons, $p_{max} = O(n)$.

The power of quantum computation [4] is mostly associated with the speed-up in computing time it can provide with respect to its classical counterpart, the paramount examples being Shor’s factoring algorithm [11] and Grover’s database search algorithm [12]. The efficiency advantage over classical computation is due essentially to the quantum superposition principle and entanglement, which allow for massively parallel information processing.

The bulk of the research effort in quantum computation has focused on the “quantization” of the classical sequential computer architecture, which has led to the quantum circuit model [4], in which information processing is realized by the sequential application of a universal set of elementary one- and two-qubit gates to typically highly entangled quantum states of many qubits. The computation is said to be efficient if the desired unitary evolution of the quantum state can be realized by the application of a polynomial number (in terms of the number of involved qubits) of these elementary quantum gates.

However, the question immediately arises if quantum mechanics can be applied successfully also to the collective information processing paradigm typical of machine intelligence algorithms and, specifically, if there are advantages in doing so. While this research has trailed the development of the quantum circuit model, it is presently experiencing a flurry of increased interest, so much so that last year NASA and Google have teamed up to found

the Quantum Artificial Intelligence Laboratory, entirely dedicated to develop and advance machine intelligence quantum algorithms.

While speed has been the main focus of quantum computation, it can be shown that quantum mechanics also offers a way out from the impossibility of reconciling the association power of content-addressable memories with the requirement of large storage capacity. Indeed, one of us pointed out already in 2001 [13–15] that storage capacity of associative memories can also be greatly enhanced by the quantum superposition principle. The key idea is to exploit the fundamental probabilistic nature of quantum mechanics. If one is willing to abandon the classical paradigm of one-off retrieval and sacrifice some speed by repeating the information retrieval step several times, then it is possible to store any desired polynomial number (in terms of the number of qubits) of patterns in a quantum associative memory and still tune the associative retrieval to a prescribed accuracy, a large advantage with respect to the classical linear limitation described above. Quantum entanglement permits to completely eliminate crosstalk and spurious memories in a tuneable probabilistic content association procedure with polynomial complexity for a polynomial number of stored patterns. Such probabilistic quantum associative memories can thus be implemented efficiently. Similar ideas in this direction were developed simultaneously in [17].

In this chapter we will review our own work on fundamental aspects of quantum associative memories and quantum pattern recognition. We will begin by a short survey of the main features of classical fully coupled neural networks like the Hopfield model and its generalizations, with a special emphasis on the capacity limitation and its origin. We will then describe the quantization of the Hopfield model [16]: the idea is to replace the classical irreversible dynamics that attracts input patterns to the closest minima of an energy function, representing the encoded memories, with a reversible unitary quantum evolution that amplifies an input quantum state to an output quantum state representing one of the stored memories at a given computational time t . In the classical model there is a complex phase diagram in terms of the two noise parameters, the temperature T and the disorder p/n with n the number of bits and p the number of stored patterns. It is, specifically the disorder due to an excessive loading factor p/n that prevents the storage of more than a critical number of patterns by causing the transition to a spin glass phase [10], even at zero temperature. Correspondingly, in the quantum version there are quantum phase transitions due to both disorder and quantum fluctuations, the latter being encoded in the effective

coupling Jt , with J being the energy parameter of the model and t being the computational time (throughout the review we will use units in which $c = 1$ and $\hbar = 1$). These are first examples of quantum collective phenomena typical of quantum machine intelligence. It turns out that, barring periodicity effects due to the unitary time evolution, the phase diagram for the quantum Hopfield model is not so different from its classical counterpart. Specifically, for small loading factors the quantum network has indeed associative power, a very interesting feature by itself, but the maximum loading factor is still limited to $p/n \leq 1$, above which there is a totally disordered spin glass phase, with no association power for any computational time. The transition to this quantum spin glass phase takes place when one tries to store a number of memories that is not anymore linearly independent.

We then turn our attention to probabilistic quantum associative memories [13–15]. The basic idea underlying their architecture is essentially the same as above, with one crucial difference: they exploit, besides a unitary evolution, a second crucial aspect of quantum mechanics, namely wave function collapse upon measurement [4]. A generic (pure) quantum state is a superposition of basis states with complex coefficients. A measurement projects (collapses) the state probabilistically onto one of the basis states, the probability distribution being governed by the squared absolute values of the superposition coefficients. Probabilistic quantum associative memories involve, besides the memory register itself a certain number b of control qubits. The unitary evolution of the input state is again determined by a Hamiltonian that depends only on the stored patterns. Contrary to quantized Hopfield memories, however, this unitary evolution mixes the memory register and the control qubits. After having applied the unitary evolution to the initial input state, the control qubits are measured. Only if one obtains a certain specific result, one proceeds to measure the memory register. This procedure is called probabilistic postselection of the measurement result and guarantees that the memory register is in a superposition of the stored patterns such that the measurement probabilities are peaked on those patterns that minimize the Hamming distance to the input. A measurement of the memory register will thus associate input and stored patterns according to this probability distribution.

Of course, if we limit ourselves to a maximum number T of repetitions, there is a non-vanishing probability that the memory retrieval will fail entirely, since the correct control qubit state will never be measured. One can say that information retrieval in these quantum memories consists of two steps: recognition (the correct state of the control qubits has been

obtained) and identification (the memory register is measured to give an output). Both steps are probabilistic and both the recognition efficiency and the identification accuracy depend on the distribution of the stored patterns: recognition efficiency is best when the number of stored patterns is large and the input is similar to a substantial cluster of them, while identification accuracy is best for isolated patterns which are very different from all other ones, both very intuitive features. Both recognition efficiency and identification accuracy can be tuned to prescribed levels by varying the repetition threshold T and the number b of control qubits.

The accuracy of the input-output association depends only on the choice of the number b of control qubits. Indeed, we will show that $t = 1/b$ plays the role of an effective temperature [15]. The lower t , the sharper is the corresponding effective Boltzmann distribution on the states closest in Hamming distance to the input and the better becomes the identification. By averaging over the distribution of stored patterns with Hamming distance to the input above a threshold d one can eliminate the dependence on the stored pattern distribution and derive the effective statistical mechanics of quantum associative memories by introducing the usual thermodynamic potentials as a function of d and the effective temperature $t = 1/b$. In particular, the free energy $F(t)$ describes the average behaviour of the recall mechanism and provides concrete criteria to tune the accuracy of the quantum associative memory. By increasing b (lowering t), the associative memory undergoes a phase transition from a disordered phase with no correlation between input and output to an ordered phase with perfect input-output association encoded in the minimal Hamming distance d . This extends to quantum information theory the relation with Ising spin systems known in error-correcting codes [18] and in public key cryptography [19].

The recognition efficiency can be tuned mainly by varying the repetition threshold T : the higher T , the larger the number of input qubits that can be corrupted without affecting recognition. The crucial point is that the recognition probability is bounded from below by $(p-1)(\pi/2)^{2b}/(pn^{2b})$. For any number of patterns, thus, a repetition threshold T polynomial in n guarantees recognition with probability $O(1)$. Due to the factor $(p-1)$ in the numerator, whose origin is exclusively quantum mechanical, the number of repetitions required for efficient recognition would actually be polynomial even for a number of patterns exponential in n . The overall complexity of probabilistic associative quantum memories is thus bounded by the complexity $O(p(2n+3))$ of the unitary evolution operator. Any polynomial number

of patterns $p = O(n^x)$ can be encoded and retrieved efficiently in polynomial computing time. The absence of spurious memories leads to a substantial storage gain with respect to classical associative memories, the price to pay being the probabilistic nature of information recall.

II. THE CLASSICAL HOPFIELD MODEL

Historically, the interest in neural networks [5] has been driven by the desire to build machines capable of performing tasks for which the traditional sequential computer architecture is not well suited, like pattern recognition, categorization and generalization. Since these higher cognitive tasks are typical of biological intelligences, the design of these parallel distributed processing systems has been largely inspired by the physiology of the human brain.

The Hopfield model is one of the best studied and most successful neural networks. It was designed to model one particular higher cognitive function of the human brain, that of associative pattern retrieval or associative memory.

The Hopfield model consists of an assembly of n binary neurons s_i , $i = 1 \dots n$ [20], which can take the values ± 1 representing their firing (+1) and resting (-1) states. The neurons are fully connected by symmetric synapses with coupling strengths $w_{ij} = w_{ji}$ ($w_{ii} = 0$). Depending on the signs of these synaptic strengths, the couplings will be excitatory (> 0) or inhibitory (< 0). The model is characterised by an energy function

$$E = -\frac{1}{2} \sum_{i \neq j} w_{ij} s_i s_j, \quad s_i = \pm 1, \quad i, j = 1 \dots n, \quad (1)$$

and its dynamical evolution is defined by the random sequential updating (in time t) of the neurons according to the rule

$$s_i(t+1) = \text{sign}(h_i(t)), \quad (2)$$

$$h_i(t) = \sum_{j \neq i} w_{ij} s_j(t), \quad (3)$$

where h_i is called the local magnetization.

The synaptic coupling strengths are chosen according to the Hebb rule

$$w_{ij} = \frac{1}{n} \sum_{\mu=1 \dots p} \xi_i^\mu \xi_j^\mu, \quad (4)$$

where ξ_i^μ , $\mu = 1 \dots p$ are p binary patterns to be memorized. An associative memory is defined as a dynamical mechanism that, upon preparing the network in an initial state s_i^0 retrieves the stored pattern ξ_i^λ that most closely resembles the presented pattern s_i^0 , where resemblance is determined by minimizing the Hamming distance, i.e. the total number of different bits in the two patterns. As emerges clearly from this definition, all the memory information in a Hopfield neural network is encoded in the synaptic strengths.

It can be easily shown that the dynamical evolution (2) of the Hopfield model satisfies exactly the requirement for an associative memory. This is because:

- The dynamical evolution (2) minimizes the energy functional (1), i.e. this energy functional never increases when the network state is updated according to the evolution rule (2). Since the energy functional is bounded by below, this implies that the network dynamics must eventually reach a stationary point corresponding to a, possibly local, minimum of the energy functional.
- The stored patterns ξ_i^μ correspond to, possibly local, minima of the energy functional. This implies that the stored patterns are attractors for the network dynamics (2). An initial pattern will evolve till it overlaps with the closest (in Hamming distance) stored pattern, after which it will not change anymore.

Actually, the second of these statements must be qualified. Indeed, the detailed behavior of the Hopfield model depends crucially upon the loading factor $\alpha = p/n$, the ratio between the number of stored memories and the number of available bits. This is best analyzed in the thermodynamic limit $p \rightarrow \infty$, $n \rightarrow \infty$, in which the different regimes can be studied by statistical mechanics techniques [5, 9] and characterized formally by the values of critical parameters.

For $\alpha < \alpha_{1\text{crit}} \simeq 0.051$, the system is in a ferromagnetic (F) phase in which there are global energy minima corresponding to all stored memories. The former differ from the original input memories only in a few percent of the total number of bits. Mixing between patterns leads to spurious local energy minima. These, however are destabilized at sufficiently high temperatures.

For $\alpha_{1\text{crit}} \simeq 0.051 < \alpha < \alpha_{2\text{crit}} \simeq 0.138$ the system is in a mixed spin glass (SG) [10] and ferromagnetic phase. There are still minima of sizable overlap with the original memories but

they are now only metastable states. The true ground state is the spin glass, characterized by an exponentially increasing number of minima due to the mixing of original memories (crosstalk). The spin glass phase is orthogonal to all stored memories. If an input pattern is sufficiently near (in Hamming distance) to one of the original memories it will be trapped by the corresponding metastable state and the retrieval procedure is successful. On the other hand, if the input pattern is not sufficiently close to one of the stored memories, the network is confused and it will end up in a state very far from all original memories.

For $\alpha > \alpha_{\text{crit}} \simeq 0.138$, the system is in a pure spin glass (*SG*) phase [10] in which all retrieval capabilities are lost due to an uncontrolled proliferation of spurious memories. It is this phase transition to a spin glass that limits the storage capacity of the Hopfield model to $\alpha = p/n < 0.138$. While various improvements are possible, the storage capacity of classical associative memories remains linearly bounded by the number n of classical bits [5].

III. QUANTUM NEURAL NETWORKS AND THE QUANTIZATION OF THE HOPFIELD MODEL

In this section we introduce a quantum information processing paradigm that is different from the standard quantum circuit model [16]. Instead of one- and two-qubit gates that are switched on and off sequentially, we will consider long-range interactions that define a fully-connected quantum neural network of qubits. This is encoded in a Hamiltonian that generates a unitary evolution in which the operator acting on one qubit depends on the collective quantum state of all the other qubits. Note that some of the most promising technologies for the implementation of quantum information processing, like optical lattices [21] and arrays of quantum dots [22] rely exactly on similar collective phenomena.

In mathematical terms, the simplest classical neural network model is a graph with the following properties:

- A state variable s_i is associated with each node (neuron) i .
- A real-valued weight w_{ij} is associated with each link (synapse) (ij) between two nodes i and j .
- A state-space-valued transfer function $f(h_i)$ of the synaptic potential $h_i = \sum_j w_{ij}s_j$ determines the dynamics of the network.

Directed graphs correspond to feed-forward neural networks [5] while undirected graphs with symmetric weights contain feed-back loops. If the graph is complete one has fully-connected neural networks like the Hopfield model. Two types of dynamical evolution have been considered: sequential or parallel synchronous. In the first case the neurons are updated one at a time according to

$$s_i(t+1) = f \left(\sum_k w_{ik} s_k(t) \right) , \quad (5)$$

while in the second case all neurons are updated at the same time. The simplest model is obtained when neurons become binary variables taking only the values $s_i = \pm 1$ for all i and the transfer function becomes the sign function. This is the original McCulloch-Pitts [20] neural network model, in which the two states represent quiescent and firing neurons.

As we have seen in the previous section, the Hopfield model [7] is a fully-connected McCulloch-Pitts network in which the synaptic weights are symmetric quantities chosen according to the Hebb rule [5]

$$w_{ij} = w_{ji} = \frac{1}{n} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu , \quad w_{ii} = 0 . \quad (6)$$

and in which the the dynamics-defining function f is the sign function, $f = \text{sign}$. This dynamics minimises the energy function

$$E = -\frac{1}{2} \sum_{i \neq j} w_{ij} s_i s_j , \quad s_i = \pm 1 , \quad i, j = 1 \dots n , \quad (7)$$

where n is the total number of neurons and ξ^μ are the p binary patterns to be memorized ($\xi_i^\mu = \pm 1$)

A quantum McCulloch-Pitts network can correspondingly be defined as a graph that satisfies:

- A two-dimensional Hilbert space \mathcal{H}_i is associated with each node (neuron) i , i.e. each neuron becomes a qubit whose basis states can be labeled as $|0\rangle$ and $|1\rangle$.
- A vector-valued weight \vec{w}_{ij} is associated with each link (synapse) (ij) between two nodes i and j .
- The synaptic potential becomes an operator $h_i = \sum_j \vec{w}_{ij} \vec{\sigma}_j$, where $\vec{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ is the vector of Pauli matrices acting on the Hilbert space \mathcal{H}_i . A unitary operator $U(h_i)$

determines the dynamics of the network starting from an initial input quantum state on the product Hilbert space of all qubits.

In case of feed-forward quantum networks on directed graphs only a subset of qubits is measured after the unitary evolution, in case of fully connected quantum networks with symmetric weights the state of the whole network is relevant.

The crucial difference with respect to classical neural networks concerns the interactions between qubits. In the classical model, the dynamics (5) induced by the transfer function is fully deterministic and irreversible, which is not compatible with quantum mechanics. A first generalization that has been considered is that of stochastic neurons, in which the transfer function determines only the probabilities that the classical state variables will take one of the two values: $n_i(t+1) = \pm 1$ with probabilities $f(\pm h_i(t))$, where f must satisfy $f(h \rightarrow -\infty) = 0$, $f(h \rightarrow +\infty) = 1$ and $f(h) + f(-h) = 1$. While this modification makes the dynamics probabilistic by introducing thermal noise, the evolution of the network is still irreversible since the actual *values* of the neurons are prescribed after an update step. In quantum mechanics the evolution must be reversible and only the magnitudes of the *changes* in the neuron variables can be postulated. Actually, the dynamics must generate a *unitary* evolution of the network.

It is known that two-level unitary gates are universal, i.e. every unitary matrix on an n -dimensional Hilbert space may be written as a product of two-level unitary matrices. However, an arbitrary unitary evolution cannot be implemented as a sequential succession of a discrete set of elementary gates, nor can it be approximated efficiently with a polynomial number of such gates [4]. In general, quantum neural networks as defined above, have to be thought of as defined by Hamiltonians H that code hard-wired qubit interactions and generate a unitary evolution $U = \exp(iHt)$. This corresponds to the parallel synchronous dynamics of classical neural networks. Only in particular cases, one of which will be the subject of the next section, does this unitary evolution admit a representation as a sequential succession of a discrete set of elementary one- and two-bit gates. In this cases the network admits a sequential dynamics as its classical counterpart.

We now describe a direct “quantization” of the Hopfield model in this spirit, i.e. by defining a quantum Hamiltonian that generalizes (7). At first sight one would be tempted to simply replace the classical spins s_i of (7) with the third Pauli matrix σ_i^z acting on the Hilbert space \mathcal{H}_i . This however would accomplish nothing, the model would still be identical

to the original classical model, since all terms in the Hamiltonian would commute between themselves. A truly quantum model must involve at least two of the three Pauli matrices. In [16] we have proposed the following “transverse” Hamiltonian:

$$\mathcal{H} = J \sum_{ij} w_{ij} \sigma_i^y \sigma_j^z, \quad (8)$$

where σ^k , $k = x, y, z$ denote the Pauli matrices and J is a coupling constant with the dimensions of mass (we remind the reader that we use units in which $c = 1, \hbar = 1$). This generates a unitary evolution of the network:

$$|\psi(t)\rangle = \exp(i\mathcal{H}t) |\psi_0\rangle, \quad (9)$$

where $|\psi_0\rangle = |\psi(t=0)\rangle$. Specifically, we will choose as initial configuration of the network the uniform superposition of all computational basis states [4]

$$|\psi_0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle. \quad (10)$$

This corresponds to a “blank memory” in the sense that all possible states have the same probability of being recovered upon measurement. In the language of spin systems this is a state in which all spins are aligned in the x direction.

Inputs ξ^{ext} can be accomodated by adding an external transverse magnetic field along the y axis, i.e. modifying the Hamiltonian to

$$\mathcal{H} = J \sum_{ij} w_{ij} \sigma_i^y \sigma_j^z + g \sum_i h_i^{\text{ext}} \sigma_i^y, \quad (11)$$

where $h_i^{\text{ext}} = \sum_j w_{ij} \xi_j^{\text{ext}}$. This external magnetic field can be thought of as arising from the interaction of the network with an additional “sensory” qubit register prepared in the state ξ^{ext} , the synaptic weights between the two layers being identical to those of the network self-couplings.

Let us now specialize to the simplest case of one assigned memory ξ in which $w_{ij} = \xi_i \xi_j / n$. In the classical Hopfield model there are two nominal stable states that represent attractors for the dynamics, the pattern ξ itself and its negative $-\xi$. Correspondingly, the quantum dynamics defined by the Hamiltonian (8) and the initial state (10) have a Z_2 symmetry generated by $\prod_i \sigma_i^x$, corresponding to the inversion $|0\rangle \leftrightarrow |1\rangle$ of all qubits.

As in the classical case we shall analyze the model in the mean field approximation. In this case, the mean field represents the average over quantum fluctuations rather than

thermal ones but the principle remains the same. The mean field model becomes exactly solvable and allows to derive self-consistency conditions on the average overlaps with the stored patterns. In the classical case, the mean field approximation is known to become exact for long-range interactions [23].

In the quantum mean-field approximation operators are decomposed in a sum of their mean values in a given quantum state and fluctuations around it, $\sigma_i^k = \langle \sigma_i^k \rangle + (\sigma_i^k - \langle \sigma_i^k \rangle)$, and quadratic terms in the fluctuations are neglected in the Hamiltonian. Apart from an irrelevant constant, this gives

$$\begin{aligned}\mathcal{H}_{\text{mf}} &= J \sum_i \sigma_i^y \left(\langle h_i^z \rangle + \frac{g}{J} h_i^{\text{ext}} \right) + \sigma_i^z \langle h_i^y \rangle , \\ \langle h_i^k \rangle &= \sum_j w_{ij} \langle \sigma_j^k \rangle = \xi_i m^k ,\end{aligned}\tag{12}$$

where $m^k = (1/n) \sum_i \langle \sigma_i^k \rangle \xi_i$ is the average overlap of the state of the network with the stored pattern. This means that each qubit i interacts with the average magnetic field (synaptic potential) $\langle h_i^k \rangle$ due to all other qubits: naturally, the correct values of these mean magnetic fields $\langle h_i^k \rangle$ have to be determined self-consistently.

To this end we compute the average pattern overlaps m^k using the mean field Hamiltonian (12) to generate the time evolution of the quantum state. This reduces to a sequence of factorized rotations in the Hilbert spaces of each qubit, giving

$$\begin{aligned}m^y &= -\frac{m^y}{|m|} \sin 2Jt|m| , \\ m^z &= \frac{m^z + (g/J)M^z}{|m|} \sin 2Jt|m| ,\end{aligned}\tag{13}$$

where $|m| = \sqrt{(m^y)^2 + (m^z + (g/J)M^z)^2}$ and $M^z = (1/n) \sum_i \xi_i^{\text{ext}} \xi_i$ is the average overlap of the external stimulus with the stored memory.

Before we present the detailed solution of these equations, let us illustrate the mechanism underlying the quantum associative memory. To this end we note that, for $g = 0$, the pattern overlaps m^y and m^z in the two directions cannot be simultaneously different from zero. As we show below, only $m^z \neq 0$ for $J > 0$ (for $J < 0$ the roles of m^y and m^z are interchanged). In this case the evolution of the network becomes a sequence of n rotations

$$\begin{pmatrix} \cos(Jt \langle h_i^z \rangle) & \sin(Jt \langle h_i^z \rangle) \\ -\sin(Jt \langle h_i^z \rangle) & \cos(Jt \langle h_i^z \rangle) \end{pmatrix}\tag{14}$$

in the two-dimensional Hilbert spaces of each qubit i . The rotation parameter is exactly the same synaptic potential h_i which governs the classical dynamics of the Hopfield model. When these rotations are applied on the initial state (10) they amount to a single update step transforming the qubit spinors into

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \cos(Jt < h_i^z >) + \sin(Jt < h_i^z >) \\ \cos(Jt < h_i^z >) - \sin(Jt < h_i^z >) \end{pmatrix}. \quad (15)$$

This is the generalization to quantum probability *amplitudes* of the probabilistic formulation of classical stochastic neurons. Indeed, the probabilities for the qubit to be in its eigenstates ± 1 after a time t , obtained by squaring the probability amplitudes, are given by $f(\pm < h^z >)$, where $f(< h^z >) = (1 + \sin(2Jt < h^z >))/2$ has exactly the properties of an activation function (alternative to the Fermi function), at least in the region $Jt < \pi/4$. In this correspondence, the effective coupling constant Jt plays the role of the inverse temperature, as usual in quantum mechanics.

We shall now focus on a network without external inputs. In this case the equation for the average pattern overlaps has only the solution $|m| = 0$ for $0 < Jt < 1/2$. For such small effective couplings (high effective temperatures), corresponding to weak synaptic connections or to short evolution times, the network is unable to remember the stored pattern. For $1/2 < Jt$, however, the solution $|m| = 0$ becomes unstable, and two new stable solutions $m^z = \pm m_0$ appear. This means that the reaction of the mean orientation of the qubit spinors against a small deviation δm^z from the $|m| = 0$ solution is larger than the deviation itself. Indeed, any so small external perturbation $(g/J)M^z$ present at the bifurcation time $t = 1/2J$ is sufficient for the network evolution to choose one of the two stable solutions, according to the sign of the external perturbation. The point $Jt = 1/2$ represents a quantum phase transition [24] from an amnesia (paramagnetic) phase to an ordered (ferromagnetic) phase in which the network has recall capabilities: the average pattern overlap m^z is the corresponding order parameter. In the ferromagnetic phase the original Z_2 symmetry of the model is spontaneously broken.

For $Jt = \pi/4$, the solution becomes $|m_0| = 1$, which means that the network is capable of perfect recall of the stored memory. For $Jt > \pi/4$ the solution m_0 decreases slowly to 0 again. Due to the periodicity of the time evolution, however, new stable solutions $m_0 = \pm 1$ appear at $Jt = (1 + 4n)\pi/4$ for every integer n . Also, for $Jt \geq 3\pi/4$, new solutions with

$m^y \neq 0$ and $m^z = 0$ appear. These, however, correspond all to metastable states. Thus, $t = \pi/4J$ is the ideal computation time for the network.

The following picture of quantum associative memories emerges from the above construction. States of the network are generic linear superpositions of computational basis states. The network is prepared in the state $|\psi_0\rangle$ and is then let to unitarily evolve for a time t . After this time the state of the network is measured, giving the result of the computation. During the evolution each qubit updates its quantum state by a rotation that depends on the aggregated synaptic potential determined by the state of all other qubits. These synaptic potentials are subject to large quantum fluctuations which are symmetric around the mean value $\langle h^z \rangle = 0$. If the interaction is strong enough, any external disturbance will cause the fluctuations to collapse onto a collective rotation of all the network's qubits towards the nearest memory.

We will now turn to the more interesting case of a finite density $\alpha = p/n$ of stored memories in the limit $n \rightarrow \infty$. In this case the state of the network can have a finite overlap with several stored memories ξ^μ simultaneously. As in the classical case we shall focus on the most interesting case of a single "condensed pattern", in which the network uniquely recalls one memory without admixtures. Without loss of generality we will chose this memory to be the first, $\mu = 1$, omitting then the memory superscript on the corresponding overlap m . Correspondingly we will consider external inputs so that only $M^{\mu=1} = M \neq 0$. For simplicity of presentation, we will focus directly on solutions with a non-vanishing pattern overlap along the z-axis, omitting also the direction superscript z .

In case of a finite density of stored patterns, one cannot neglect the noise effect due to the infinite number of memories. This changes (13) to

$$\begin{aligned} m &= \frac{1}{n} \sum_i \sin 2Jt \left(m + \frac{g}{J} M + \Delta_i \right) , \\ \Delta_i &= \sum_{\mu \neq 1} \xi_i^1 \xi_i^\mu m^\mu . \end{aligned} \tag{16}$$

As in the classical case we will assume that $\{\xi_i^\mu\}$ and $\{m^\mu, \mu \neq 1\}$ are all independent random variables with mean zero and we will denote by square brackets the configurational average over the distributions of these random variables. As a consequence of this assumption, the mean and variance of the noise term are given by $[\Delta_i] = 0$ and $[\Delta_i^2] = \alpha r$, where

$$r = \frac{1}{\alpha} \sum_{\mu \neq 1} [(m^\mu)^2] \tag{17}$$

is the spin-glass order parameter [10]. According to the central limit theorem one can now replace $n^{-1} \sum_i$ in (16) by an average over a Gaussian noise,

$$m = \int \frac{dz}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \sin 2Jt \left(m + \frac{g}{J}M + \sqrt{\alpha r}z \right) . \quad (18)$$

The second order parameter r has to be evaluated self-consistently by a similar procedure starting from the equation analogous to eq. (16) for $\mu \neq 1$. In this case one can use $m^\mu \ll 1$ for $\mu \neq 1$ to expand the transcendental function on the right-hand side in powers of this small parameter, which gives

$$\begin{aligned} v &= \int \frac{dz}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \sin^2 2Jt \left(m + \frac{g}{J}M + \sqrt{\alpha r}z \right) , \\ x &= \int \frac{dz}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \cos 2Jt \left(m + \frac{g}{J}M + \sqrt{\alpha r}z \right) , \end{aligned} \quad (19)$$

where $v = (1 - 2Jtx)^2 r$. Solving the integrals gives finally the following coupled equations for the two order parameters m and r :

$$\begin{aligned} m &= \sin 2Jt \left(m + \frac{g}{J}M \right) e^{-2(Jt)^2 \alpha r} , \\ r &= \frac{1}{2} \frac{1 - \cos 4Jt \left(m + \frac{g}{J}M \right) e^{-8(Jt)^2 \alpha r}}{\left(1 - 2Jt \cos 2Jt \left(m + \frac{g}{J}M \right) e^{-2(Jt)^2 \alpha r} \right)^2} . \end{aligned} \quad (20)$$

In terms of these order parameters one can distinguish three phases of the network. First of all the value of m determines the presence ($m > 0$) or absence ($m = 0$) of ferromagnetic order (F). If $m = 0$ the network can be in a paramagnetic phase (P) if also $r = 0$ or a quantum spin glass phase (SG) if $r > 0$. The phase structure resulting from a numerical solution of the coupled equations (20) for $g = 0$ is shown in Fig. 1.

For $\alpha < 0.025$ the picture is not very different from the single memory case. For large enough computation times there exists a ferromagnetic phase in which the $m = 0$ solution is unstable and the network has recall capabilities. The only difference is that the maximum value of the order parameter m is smaller than 1 (recall is not perfect due to noise) and the ideal computation time t at which the maximum is reached depends on α . For $0.025 < \alpha < 1.000$ instead, ferromagnetic order coexists as a metastable state with a quantum spin glass state. This means that ending up in the memory retrieval solution depends not only on the presence of an external stimulus but also on its magnitude; in other words, the external pattern has to be close enough to the stored memory in order to be retrieved. For $1 < \alpha$ all retrieval capabilities are lost and the network will be in a quantum spin glass state for all

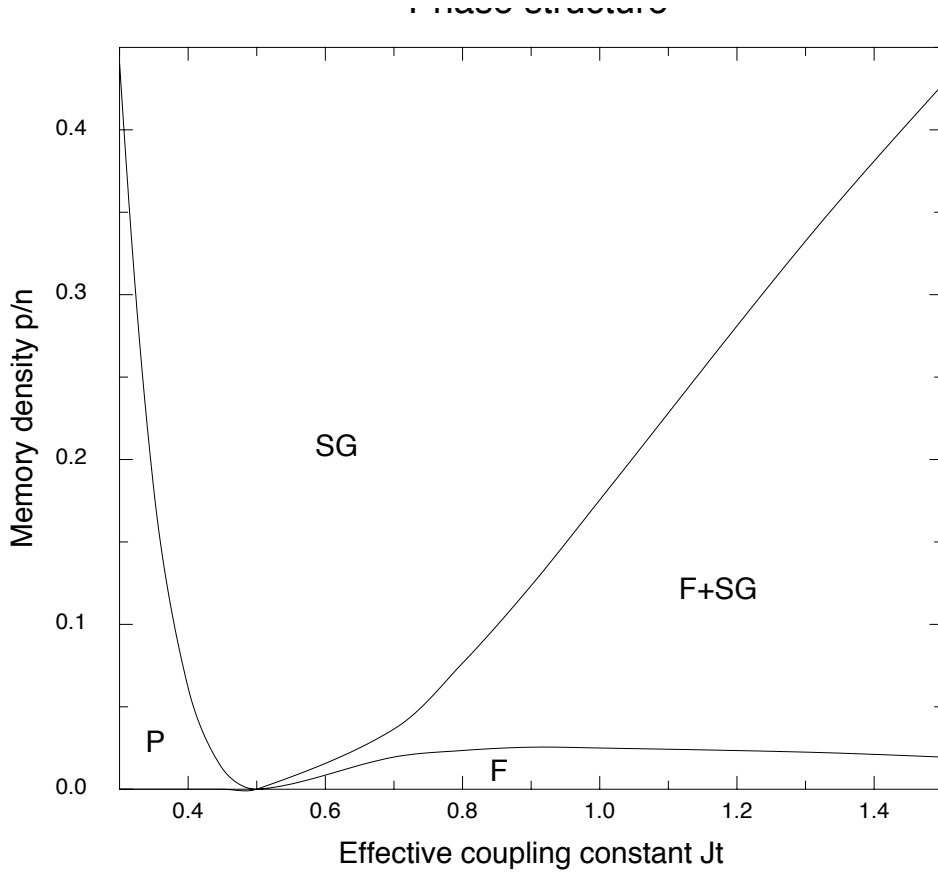


FIG. 1: The phase structure of quantum associative memories with finite density of stored patterns. P, F and SG denote (quantum) paramagnetic, ferromagnetic and spin-glass phases, respectively. F + SG denotes a mixed phase in which the memory retrieval solution is only locally stable.

computation times (after the transition from the quantum paramagnet). $\alpha = 1$ is thus the maximum memory capacity of this quantum network. Note that $\alpha = 1$ corresponds to the maximum possible number of linearly independent memories. For memory densities smaller but close to this maximum value, however, the ferromagnetic solution exists only for a small range of effective couplings centered around $Jt \simeq 9$: for these high values of Jt the quality of pattern retrieval is poor, the value of the order parameter m being of the order 0.15-0.2. Much better retrieval qualities are obtained for smaller effective couplings: e.g. for $Jt = 1$ the order parameter is larger than 0.9 (corresponding to an error rate smaller than 5%) for memory densities up to 0.1. In this case, however the maximum memory density is 0.175,

comparable with the classical result of the Hopfield model. Quantum mechanics, here, does not carry any advantage.

IV. PROBABILISTIC QUANTUM MEMORIES

We have seen in the last section that crosstalk prevents the amplification of patterns stored in the weights of a simple quantum Hamiltonian like (8) when the loading factor exceeds a linear bound comparable with the classical one. In this section we show that this limit can be overcome by probabilistic quantum memories, which use postselection of the measurement results of certain control qubits [13–15]. The price to pay is that such probabilistic memories require repetitions of the retrieval process and that there is non-vanishing probability that this fails entirely. When it is successful, however, it allows retrieval of the most appropriate pattern among a polynomial pool instead of a linear one.

A. Storing patterns

Let us start by describing the elementary quantum gates [4] that we will use in the rest of the paper. First of all there are the single-qbit gates represented by the Pauli matrices σ^i , $i = x, y, z$. The first Pauli matrix σ^x , in particular, implements the NOT gate. Another single-qbit gate is the Hadamard gate H , with the matrix representation

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (21)$$

Then, we will use extensively the two-qbit XOR (exclusive OR) gate, which performs a NOT on the second qbit if and only if the first one is in state $|1\rangle$. In matrix notation this gate is represented as $\text{XOR} = \text{diag}(1, \sigma^x)$, where 1 denotes a two-dimensional identity matrix and σ^x acts on the components $|01\rangle$ and $|11\rangle$ of the Hilbert space. The 2XOR, or Toffoli gate is the three qbit generalization of the XOR gate: it performs a NOT on the third qbit if and only if the first two are both in state $|1\rangle$. In matrix notation it is given by $2\text{XOR} = \text{diag}(1, 1, \sigma^x)$. In the storage algorithm we shall make use also of the nXOR generalization of these gates, in which there are n control qbits. This gate is also used in the subroutines implementing the oracles underlying Grover's algorithm [4] and can be realized using unitary maps affecting only few qbits at a time [25], which makes it efficient. All these

are standard gates. In addition to them we introduce the two-qbit controlled gates

$$\begin{aligned} CS^i &= |0\rangle\langle 0| \otimes 1 + |1\rangle\langle 1| \otimes S^i, \\ S^i &= \begin{pmatrix} \sqrt{\frac{i-1}{i}} & \frac{1}{\sqrt{i}} \\ -\frac{1}{\sqrt{i}} & \sqrt{\frac{i-1}{i}} \end{pmatrix}, \end{aligned} \quad (22)$$

for $i = 1, \dots, p$. These have the matrix notation $CS^i = \text{diag}(1, S^i)$. For all these gates we shall indicate by subscripts the qbits on which they are applied, the control qbits coming always first.

The construction of quantum memories relies, of course, on the fundamental fact that one can use entanglement to "store" an arbitrary number p of binary patterns p^i of length n in a quantum superposition of just n qubits,

$$|m\rangle = \frac{1}{\sqrt{p}} \sum_{i=1}^p |p^i\rangle. \quad (23)$$

The idea of the memory architecture consists thus of two steps:

- Generate the state $|m\rangle$ by a unitary evolution M from a simple prepared state, say $|0_1, \dots, 0_n\rangle$, $|m\rangle = M|0_1, \dots, 0_n\rangle$.
- Given an input state $|i\rangle = |i_1, \dots, i_n\rangle$, generate from $|m\rangle$ a superposition of the pattern states that is no more uniform but whose amplitudes define a probability distribution peaked on the pattern states with minimal Hamming distance front the input. It is this step that involves both a unitary evolution and a postselection of the measurement result.

The quantum memory itself is the unitary operator M that codes the p patterns. It defines implicitly a Hamiltonian through the formal relation $M = \exp(i\mathcal{H})$, a Hamiltonian that represents pattern-dependent interactions among the qubits. This is the quantum generalization of the classical Hopfield model. In order to dispel any possible misunderstandings right away, we point out that this is quite different to the communication of classical information via a quantum channel, limited by the Holevo theorem [26], as we discuss in detail below.

In order to construct explicitly the quantum memory M we will start from an algorithm that loads sequentially the classical patterns into an auxiliary register, from which they

are then copied into the actual memory register. A first version of such an algorithm was introduced in [27]. The simplified version that we present here is due to [13].

We shall use three registers: a first register p of n qbits in which we will subsequently feed the patterns p^i to be stored, a utility register u of two qbits prepared in state $|01\rangle$, and another register m of n qbits to hold the memory. This latter will be initially prepared in state $|0_1, \dots, 0_n\rangle$. The full initial quantum state is thus

$$|\psi_0^1\rangle = |p_1^1, \dots, p_n^1; 01; 0_1, \dots, 0_n\rangle . \quad (24)$$

The idea of the storage algorithm is to separate this state into two terms, one corresponding to the already stored patterns, and another ready to process a new pattern. These two parts will be distinguished by the state of the second utility qbit u_2 : $|0\rangle$ for the stored patterns and $|1\rangle$ for the processing term.

For each pattern p^i to be stored one has to perform the operations described below:

$$|\psi_1^i\rangle = \prod_{j=1}^n 2XOR_{p_j^i u_2 m_j} |\psi_0^i\rangle . \quad (25)$$

This simply copies pattern p^i into the memory register of the processing term, identified by $|u_2\rangle = |1\rangle$.

$$\begin{aligned} |\psi_2^i\rangle &= \prod_{j=1}^n NOT_{m_j} XOR_{p_j^i m_j} |\psi_1^i\rangle , \\ |\psi_3^i\rangle &= nXOR_{m_1 \dots m_n u_1} |\psi_2^i\rangle . \end{aligned} \quad (26)$$

The first of these operations makes all qbits of the memory register $|1\rangle$'s when the contents of the pattern and memory registers are identical, which is exactly the case only for the processing term. Together, these two operations change the first utility qbit u_1 of the processing term to a $|1\rangle$, leaving it unchanged for the stored patterns term.

$$|\psi_4^i\rangle = CS_{u_1 u_2}^{p+1-i} |\psi_3^i\rangle . \quad (27)$$

This is the central operation of the storing algorithm. It separates out the new pattern to be stored, already with the correct normalization factor.

$$\begin{aligned} |\psi_5^i\rangle &= nXOR_{m_1 \dots m_n u_1} |\psi_4^i\rangle , \\ |\psi_6^i\rangle &= \prod_{j=n}^1 XOR_{p_j^i m_j} NOT_{m_j} |\psi_5^i\rangle . \end{aligned} \quad (28)$$

These two operations are the inverse of eqs.(26) and restore the utility qbit u_1 and the memory register m to their original values. After these operations on has

$$|\psi_6^i\rangle = \frac{1}{\sqrt{p}} \sum_{k=1}^i |p^i; 00; p^k\rangle + \sqrt{\frac{p-i}{p}} |p^i; 01; p^i\rangle . \quad (29)$$

With the last operation,

$$|\psi_7^i\rangle = \prod_{j=n}^1 2XOR_{p_j^i u_2 m_j} |\psi_6^i\rangle , \quad (30)$$

one restores the third register m of the processing term, the second term in eq.(29) above, to its initial value $|0_1, \dots, 0_n\rangle$. At this point one can load a new pattern into register p and go through the same routine as just described. At the end of the whole process, the m -register is exactly in state $|m\rangle$, eq. (23).

Any quantum state can be generically obtained by a unitary transformation of the initial state $|0, \dots, 0\rangle$. This is true also for the memory state $|m\rangle$. In the following we will explicitly construct the unitary memory operator M which implements the transformation $|m\rangle = M |0, \dots, 0\rangle$.

To this end we introduce first the single-qbit unitary gates

$$U_j^i = \cos\left(\frac{\pi}{2} p_j^i\right) 1 + i \sin\left(\frac{\pi}{2} p_j^i\right) \sigma_2 , \quad (31)$$

where σ_2 is the second Pauli matrix. These operators are such that their product over the n qbits generates pattern p^i out of $|0, \dots, 0\rangle$:

$$\begin{aligned} |p^i\rangle &= P^i |0, \dots, 0\rangle , \\ P^i &\equiv \prod_{j=1}^n U_j^i . \end{aligned} \quad (32)$$

We now introduce, in addition to the memory register proper, the same two utility qbits as before, also initially in the state $|0\rangle$. The idea is, exactly as in the sequential algorithm, to split the state into two parts, a storage term with $|u_2\rangle = |0\rangle$ and a processing term with $|u_2\rangle = |1\rangle$. Therefore we generalize the operators P^i defined above to

$$CP_{u_2}^i \equiv \prod_{j=1}^n CU_{u_2 j}^i , \quad (33)$$

which loads pattern p^i into the memory register only for the processing term. It is then easy

to check that

$$\begin{aligned}
|m; 00\rangle &= M |0, \dots, 0; 00\rangle, \\
M &= \prod_{i=1}^p \left[(CP_{u_2}^i)^{-1} NOT_{u_1} CS_{u_1 u_2}^{p+1-i} XOR_{u_2 u_1} CP_{u_2}^i \right] \times \\
&\times NOT_{u_2}.
\end{aligned} \tag{34}$$

From this construction it is easy to see that the memory operator M involves a number $p(2n + 3) + 1$ of elementary one- and two-qbit gates. It is thus efficient for any number p of patterns polynomial in the number n of qubits. It is interesting to note that another version of this operator has been recently derived in [28], with a bound of $O(pn^3/6)$ on its complexity. This is also linear in p , implying again efficiency for a polynomial number of patterns.

While the memory construction we have presented here mirrors its classical counterpart, it is important to stress one notable difference. In classical associative memories, patterns are stored as minima of an energy landscape or, alternatively in the parameters of a dynamical evolution law [5]. This is reflected verbatim in the construction of the unitary operator M in (34), which completely codes the patterns in a dynamical law, albeit reversible in the quantum case. In quantum mechanics, however, there is the possibility of shuffling some (but not all, as we will shortly see) information about the patterns from the unitary evolution law M onto a set of quantum states.

The ideal, most compressed quantum memory would indeed be the quantum superposition of patterns $|m\rangle$ in (23) itself. This, however is impossible. If the memory state has to be used for information retrieval it must be measured and this destroys all information about the patterns (save the one obtained in the measurement). The quantum state must therefore be copied prior to use and this is impossible since the linearity of quantum mechanics forbids exact universal cloning of quantum states [29]. Universal cloning of quantum states is possible only in an approximate sense [30] and has two disadvantages: first of all the copies are imperfect, though optimal [31] and secondly, the quality of the master copy decreases with each additional copy made. Approximate universal cloning is thus excluded for the purposes of information recall since the memory would be quickly washed out.

This leaves state-dependent cloning [32] as the only viable option. State-dependent cloners are designed to reproduce only a finite number of states and this is definitely enough for our purposes. Actually the memory M in (34) is equivalent to a state-dependent cloner

for the the state $|m\rangle$ in (23). In this case the information about the stored patterns is completely coded in the memory operator, or equivalently, the state-dependent cloner. It is possible, however, to subdivide the pattern information among an operator and a set of quantum states, obviously including $|m\rangle$, by using a probabilistic cloning machine [33]. Probabilistic cloners copy quantum states exactly but the copying process is not guaranteed to succeed and must be repeated until the measurement of an auxiliary register produces a given result associated with copying success. In general, any number of linearly independent states can be copied probabilistically. In the present case for example, it would be sufficient to consider any dummy state $|d\rangle$ different from $|m\rangle$ (for more than two states the condition would be linear independence) and to construct a probabilistic cloning machine for these two states. This machine would reproduce $|m\rangle$ with probability p_m and $|d\rangle$ with probability p_d ; a flag would tell when the desired state $|m\rangle$ has been obtained. In order to obtain an exact copy of $|m\rangle$ one would need then $1/p_m$ trials on average. The master copy would be exactly preserved.

The cloning efficiencies of the probabilistic cloner of two states are bounded as follows [33]:

$$p_m + p_d \leq \frac{2}{1 + \langle d|m \rangle} . \quad (35)$$

This bound can be made large by choosing $|d\rangle$ as nearly orthogonal to $|m\rangle$ as possible. A simple way to achieve this for a large number of patterns would be, for example, to encode also the state

$$|d\rangle = \frac{1}{\sqrt{p}} \sum_{i=1}^p (-1)^{i+1} |p^i\rangle \quad (36)$$

together with $|m\rangle$ when storing information. This can be done easily by using alternately the operators S^i and $(S^i)^{-1}$ in the storing algorithm above. For binary patterns which are all different from one would then have

$$\begin{aligned} \langle d|m \rangle &= 0 , & p \text{ even} , \\ \langle d|m \rangle &= \frac{1}{p} , & p \text{ odd} , \end{aligned} \quad (37)$$

and the bound for the cloning efficiencies would be very close to its maximal value 2 in both cases.

The quantum network for the probabilistic cloner of two states has been developed in [34]. It can be constructed exclusively out of the two simple distinguishability transfer (D)

and state separation (S) gates. As expected, these gates embody information about the two states to be cloned. Part of the memory, therefore, still resides in the cloning network. The pattern-dependence of the network cloner can be decreased by choosing a larger set of states in the pool that can be cloned, so that the cloner becomes more and more generic. On one side this decreases also the efficiency of the cloner, so that more repetitions are required, on the other side, since the clonable pool is limited to a set of linearly independent states, one can never eliminate completely the pattern-dependence of the cloning operator. This is why the original claim of an exponential capacity increase of quantum associative memories [13], based on probabilistic cloning of the state $|m\rangle$, is excessive. The complexity of the cloner, be it exact as in the memory operator M or probabilistic, remains linear in the number of patterns and the requirement of efficient implementability limits thus p to a polynomial function of the number n of qubits., which is still a large improvement upon classical associative memories.

B. Retrieving patterns

Let us now assume we are given a binary input i that is a corrupted version of one of the patterns stored in the memory. The task of the retrieval algorithm is to "recognize" it, i.e. output the stored pattern that most resembles this input, where similarity is defined (here) in terms of the Hamming distance, the number of different bits between the two patterns, although other similarity measures [?] could also be incorporated.

The retrieval algorithm requires also three registers. The first register i of n qubits contains the input pattern; the second register m , also of n qubits, contains the memory $|m\rangle$; finally there is a control register c with b qubits all initialized in the state $|0\rangle$. The full initial quantum state is thus:

$$|\psi_0\rangle = \frac{1}{\sqrt{p}} \sum_{k=1}^p |i; p^k; 0_1, \dots, 0_b\rangle \quad (38)$$

where $|i\rangle = |i_1, \dots, i_n\rangle$ denotes the input qubits, the second register, m , contains the memory (23) and all b control qubits are in state $|0\rangle$. Applying the Hadamard gate to the first control

qbit one obtains

$$\begin{aligned} |\psi_1\rangle &= \frac{1}{\sqrt{2p}} \sum_{k=1}^p |i; p^k; 0_1, \dots, 0_b\rangle \\ &+ \frac{1}{\sqrt{2p}} \sum_{k=1}^p |i; p^k; 1_1, \dots, 0_b\rangle . \end{aligned} \quad (39)$$

Let us now apply to this state the following combination of quantum gates:

$$|\psi_2\rangle = \prod_{j=1}^n NOT_{m_j} XOR_{i_j m_j} |\psi_1\rangle , \quad (40)$$

As a result of the above operation the memory register qbits are in state $|1\rangle$ if i_j and p_j^k are identical and $|0\rangle$ otherwise:

$$\begin{aligned} |\psi_2\rangle &= \frac{1}{\sqrt{2p}} \sum_{k=1}^p |i; d^k; 0_1, \dots, 0_b\rangle \\ &+ \frac{1}{\sqrt{2p}} \sum_{k=1}^p |i; d^k; 1_1, \dots, 0_b\rangle , \end{aligned} \quad (41)$$

where $d_j^k = 1$ if and only if $i_j = p_j^k$ and $d_j^k = 0$ otherwise.

Consider now the following Hamiltonian:

$$\begin{aligned} \mathcal{H} &= (d_H)_m \otimes (\sigma^z)_{c_1} , \\ (d_H)_m &= \sum_{j=1}^n \left(\frac{\sigma^z + 1}{2} \right)_{m_j} , \end{aligned} \quad (42)$$

where σ^z is the third Pauli matrix. \mathcal{H} measures the number of 0's in register m , with a plus sign if c_1 is in state $|0\rangle$ and a minus sign if c_1 is in state $|1\rangle$. Given how we have prepared the state $|\psi_2\rangle$, this is nothing else than the number of qbits which are different in the input and memory registers i and m . This quantity is called the *Hamming distance* and represents the (squared) Euclidean distance between two binary patterns.

Every term in the superposition (41) is an eigenstate of \mathcal{H} with a different eigenvalue. Applying thus the unitary operator $\exp(i\pi\mathcal{H}/2n)$ to $|\psi_2\rangle$ one obtains

$$\begin{aligned} |\psi_3\rangle &= e^{i\frac{\pi}{2n}\mathcal{H}} |\psi_2\rangle , \\ |\psi_3\rangle &= \frac{1}{\sqrt{2p}} \sum_{k=1}^p e^{i\frac{\pi}{2n}d_H(i,p^k)} |i; d^k; 0_1, \dots, 0_b\rangle \\ &+ \frac{1}{\sqrt{2p}} \sum_{k=1}^p e^{-i\frac{\pi}{2n}d_H(i,p^k)} |i; d^k; 1_1, \dots, 0_b\rangle , \end{aligned} \quad (43)$$

where $d_H(i, p^k)$ denotes the Hamming distance between the input i and the stored pattern p^k .

In the final step we restore the memory gate to the state $|m\rangle$ by applying the inverse transformation to eq. (40) and we apply the Hadamard gate to the control qbit c_1 , thereby obtaining

$$\begin{aligned} |\psi_4\rangle &= H_{c_1} \prod_{j=n}^1 \text{XOR}_{i_j m_j} \text{NOT}_{m_j} |\psi_3\rangle, \\ |\psi_4\rangle &= \frac{1}{\sqrt{p}} \sum_{k=1}^p \cos \frac{\pi}{2n} d_H(i, p^k) |i; p^k; 0_1, \dots, 0_b\rangle \\ &\quad + \frac{1}{\sqrt{p}} \sum_{k=1}^p \sin \frac{\pi}{2n} d_H(i, p^k) |i; p^k; 1_1, \dots, 0_b\rangle. \end{aligned} \quad (44)$$

The idea is now to repeat the above operations sequentially for all b control qbits c_1 to c_b . This gives

$$\begin{aligned} |\psi_{\text{fin}}\rangle &= \frac{1}{\sqrt{p}} \sum_{k=1}^p \sum_{l=0}^b \cos^{b-l} \left(\frac{\pi}{2n} d_H(i, p^k) \right) \times \\ &\quad \sin^l \left(\frac{\pi}{2n} d_H(i, p^k) \right) \sum_{\{J^l\}} |i; p^k; J^l\rangle, \end{aligned} \quad (45)$$

where $\{J^l\}$ denotes the set of all binary numbers of b bits with exactly l bits 1 and $(b-l)$ bits 0.

Note that one could also dispense with a register for the input but, rather, code also the input directly into a unitary operator. Indeed, the auxiliary quantum register for the input is needed only by the operator (40) leading from (39) to (41). The same result (apart from an irrelevant overall sign) can be obtained by applying

$$\begin{aligned} I &= \prod_{j=1}^n U_j, \\ U_j &= \sin \left(\frac{\pi}{2} i_j \right) 1 + i \cos \left(\frac{\pi}{2} i_j \right) \sigma_2, \end{aligned} \quad (46)$$

directly on the memory state $|m\rangle$. The rest of the algorithm is the same, apart the reversing of the operator (40) which needs now the operator I^{-1} .

The end effect of the information retrieval algorithm represents thus a rotation of the memory quantum state in the enlarged Hilbert space obtained by adding b control qbits. The overall effect of this rotation is an amplitude concentration on memory states similar to

the input, if there is a large number of $|0\rangle$ control qbits in the output state and an amplitude concentration on states different from the input, if there is a large number of $|1\rangle$ control qbits in the output state. As a consequence, the most interesting state for information retrieval purposes is the projection of $|\psi_{\text{fin}}\rangle$ onto the subspace with all control qbits in state $|0\rangle$.

There are two ways of obtaining this projection. The first, and easiest one, is to simply repeat the above algorithm and measure the control register several times, until exactly the desired state for the control register is obtained. If the number of such repetitions exceeds a preset threshold T the input is classified as "non-recognized" and the algorithm is stopped. Otherwise, once $|c_1, \dots, c_b\rangle = |0_1, \dots, 0_b\rangle$ is obtained, one proceeds to a measurement of the memory register m , which yields the output pattern of the memory.

The second method is to first apply T steps of the amplitude amplification algorithm [35] rotating $|\psi_{\text{fin}}\rangle$ towards its projection onto the "good" subspace formed by the states with all control qbits in state $|0\rangle$. To this end it is best to use the version of the retrieving algorithm that does not need an auxiliary register for the input. Let us define as $R(i)$ the input-dependent operator which rotates the memory state in the Hilbert space enlarged by the b control qbits towards the final state $|\psi_{\text{fin}}\rangle$ in eq. (45) (where we now omit the auxiliary register for the input):

$$|\psi_{\text{fin}}\rangle = R(i) |m; 0_1, \dots, 0_b\rangle . \quad (47)$$

By adding also the two utility qbits needed for the storing algorithm one can then obtain $|\psi_{\text{fin}}\rangle$ as a unitary transformation of the initial state with all qbits in state $|0\rangle$:

$$|\psi_{\text{fin}}; 00\rangle = R(i)M |0, \dots, 0; 0_1, \dots, 0_b; 00\rangle . \quad (48)$$

The amplitude amplification rotation of $|\psi_{\text{fin}}; 00\rangle$ towards its "good" subspace in which all b control qbits are in state $|0\rangle$ is then obtained [35] by repeated application of the operator

$$Q = -R(i)MS_0M^{-1}R^{-1}(i)S \quad (49)$$

on the state $|\psi_{\text{fin}}; 00\rangle$. Here S conditionally changes the sign of the amplitude of the "good" states with the b control qbits in state $|0\rangle$, while S_0 changes the sign of the amplitude if and only if the state is the zero state $|0, \dots, 0; 0_1, \dots, 0_b; 00\rangle$. As before, if a measurement of the control register after the T iterations of the amplitude amplification rotation yields $|0_1, \dots, 0_b\rangle$ one proceeds to a measurement of the memory register, otherwise the input is classified as "non-recognized".

The expected number of repetitions needed to measure the desired control register state is $1/P_b^{\text{rec}}$, with

$$P_b^{\text{rec}} = \frac{1}{p} \sum_{k=1}^p \cos^{2b} \left(\frac{\pi}{2n} d_H(i; p^k) \right) \quad (50)$$

the probability of measuring $|c_1, \dots, c_n\rangle = |0_1, \dots, 0_n\rangle$. The threshold T governs thus the *recognition efficiency* of the memory. Note, however, that amplitude amplification provides a quadratic boost [35] to the recognition efficiency since only $1/\sqrt{P_b^{\text{rec}}}$ steps are typically required to rotate $|\psi_{\text{fin}}\rangle$ onto the desired subspace. Accordingly, the threshold T can be lowered to \sqrt{T} with respect to the method of projection by measurement. The crucial point is that, due to the quantum nature of the retrieval mechanism, this recognition probability depends on the distribution of *all* stored patterns. A lower bound on the recognition probability can thus be established as follows. Of all the stored patterns, all but one have Hamming distance from the input smaller or equal than $(n-1)$. There is only pattern that can have a larger Hamming distance equal to n . So we shall use the upper bound $(n-1)$ for the Hamming distance of all patterns but one, for which we shall use the upper bound n , and this one does not contribute to the recognition probability since the cosine function vanishes. Given that cosine is a decreasing function in the interval $[0, \pi/2]$, we get the lower bound

$$P_b^{\text{rec}} \geq P_b^{\text{min}} = \frac{p-1}{p} \cos^{2b} \left(\frac{\pi(n-1)}{2n} \right) . \quad (51)$$

For $n \gg 1$ we can now estimate this lower bound as

$$P_b^{\text{min}} \simeq \frac{p-1}{p} \left(\frac{\pi}{2n} \right)^{2b} . \quad (52)$$

This shows that, independent of the number p of patterns, the threshold T for recognition can be set as a polynomial function of the number n of qubits. Note that this is entirely due to the factor $(p-1)$ in the numerator of (52), which, in turn, depends on the quantum nature of the memory. In other words, the probabilistic character of the retrieval process does not limit at all the number of possible stored patterns, the typical number of repetitions required would be polynomial even for an exponential number of patterns. The efficient implementability of the quantum memory is limited only by the number of elementary quantum gates in M , which is linear in p .

In general, the probability of recognition is determined by comparing (even) powers of cosines and sines of the distances to the stored patterns. It is thus clear that the worst case

for recognition is the situation in which there is an isolated pattern, with the remaining patterns forming a tight cluster spanning all the largest distances to the first one. As a consequence, the threshold needed to recognize all patterns diminishes when the number of stored patterns becomes very large, since, in this case, the distribution of patterns becomes necessarily more homogeneous. Indeed, for the maximal number of stored patterns $p = 2^n$ one has $P_b^{\text{rec}} = 1/2^b$ and the recognition efficiency becomes also maximal, as it should be.

Once the input pattern i is recognized, the measurement of the memory register yields the stored pattern p^k with probability

$$P_b(p^k) = \frac{1}{Z} \cos^{2b} \left(\frac{\pi}{2n} d_H(i, p^k) \right) , \quad (53)$$

$$Z = p P_b^{\text{rec}} = \sum_{k=1}^p \cos^{2b} \left(\frac{\pi}{2n} d_H(i, p^k) \right) . \quad (54)$$

Clearly, this probability is peaked around those patterns which have the smallest Hamming distance to the input. The highest probability of retrieval is thus realized for that pattern which is most similar to the input. This is always true, independently of the number of stored patterns. In particular, contrary to classical associative memories, there are *no spurious memories*: the probability of obtaining as output a non-stored pattern is always zero. This is another manifestation of the fact that there are no restrictions on the loading factor p/n due to the information retrieval algorithm.

In addition to the threshold T , there is a second tunable parameter, namely the number b of control qbits. This new parameter b controls the *identification efficiency* of the quantum memory since, increasing b , the probability distribution $P_b(p^k)$ becomes more and more peaked on the low $d_H(i, p^k)$ states, until

$$\lim_{b \rightarrow \infty} P_b(p^k) = \delta_{kk_{\min}} , \quad (55)$$

where k_{\min} is the index of the pattern (assumed unique for convenience) with the smallest Hamming distance to the input.

While the recognition efficiency depends on comparing powers of cosines and sines of the same distances in the distribution, the identification efficiency depends on comparing the (even) powers of cosines of the different distances in the distribution. Specifically, it is best when one of the distances is zero, while all others are as large as possible, such that the probability of retrieval is completely peaked on one pattern. As a consequence, the identification efficiency is best when the recognition efficiency is worst and viceversa.

The role of the parameter b becomes familiar upon a closer examination of eq.(53). Indeed, the quantum distribution described by this equation is equivalent to a canonical Boltzmann distribution with (dimensionless) temperature $t = 1/b$ and (dimensionless) energy levels

$$E^k = -2 \log \cos \left(\frac{\pi}{2n} d_H(i, p^k) \right) , \quad (56)$$

with Z playing the role of the partition function.

The appearance of an effective thermal distribution suggests studying the average behaviour of quantum associative memories via the corresponding thermodynamic potentials. Before this can be done, however, one must deal with the different distributions of stored patterns characterizing each individual memory. The standard way to do this in similar classical problems is to average over the random distribution of patterns. Typically, one considers quenched averages in which extensive quantities, like the free energy are averaged over the disorder: this is the famed replica trick used to analyze spin glasses [10]. In the present case, however, the disorder cannot lead to spin-glass-like phases since there are no spurious memories: by construction, probabilistic quantum memories can output only one of the stored patterns. The only question is how accurate is the retrieval of the most similar pattern to the input as a function of the fictitious temperature $t = 1/b$. To address this question we will "quench" only one aspect of the random pattern distribution, namely the minimal Hamming distance d between the input and the stored patterns. The rest of the random pattern distribution will be considered as annealed. In doing so, one obtains an average description of the average memory as a function of the fictitious temperature $t = 1/b$ and the minimal Hamming distance d .

To do so we first normalize the pattern representation by adding (modulo 2) to all patterns, input included, the input pattern i . This clearly preserves all Hamming distances and has the effect of choosing the input as the state with all qbits in state $|0\rangle$. The Hamming distance $d_H(i, p^k)$ becomes thus simply the number of qbits in pattern p^k with value $|1\rangle$. The averaged partition function takes then a particularly simple form:

$$Z_{\text{av}} = \frac{p}{N_\lambda} \sum_{\{\lambda\}} \sum_{j=d}^n \lambda_j \cos^{2b} \left(\frac{\pi}{2} \frac{j}{n} \right) , \quad (57)$$

where λ_j describes a probability distribution, $\sum_{j=d}^n \lambda_j = 1$, with the following properties. Let the number of patterns scale as the x^{th} power of the number of qubits, $p = \alpha_x n^x$ for

$n \gg 1$. Then

$$\begin{aligned}\lambda_j &= 0, & j &> n - x, \\ \lambda_j &\leq \frac{1}{\alpha_x x!}, & j &= n - x,\end{aligned}\tag{58}$$

with all other λ_j for $j < n - x$ unconstrained. $\{\lambda\}$ is the set of such distributions and N_λ the corresponding normalization factor. Essentially the probability distribution becomes unconstrained in the limit of large n .

We now introduce the free energy $F(b, d)$ by the usual definition

$$Z_{\text{av}} = p e^{-bF(b, d)} = Z_{\text{av}}(b = 0) e^{-bF(b, d)},\tag{59}$$

where we have chosen a normalization such that $\exp(-bF)$ describes the deviation of the partition function from its value for $b = 0$ (high effective temperature). Since Z/p , and consequently also Z_{av}/p posses a finite, non-vanishing large- n limit, this normalization ensures that $F(b, d)$ is intensive, exactly like the energy levels (56), and scales as a constant for large n . This is the only difference with respect to the familiar situation in statistical mechanics.

The free energy describes the equilibrium of the system at effective temperature $t = 1/b$ and has the usual expression in terms of the internal energy U and the entropy S :

$$\begin{aligned}F(t, d) &= U(t, d) - tS(t, d), \\ U(t, d) &= \langle E \rangle_t, \quad S(t, d) = \frac{-\partial F(t, d)}{\partial t}.\end{aligned}\tag{60}$$

Note that, with the normalization we have chosen in (59), the entropy S is always a negative quantity describing the deviation from its maximal value $S_{\text{max}} = 0$ at $t = \infty$.

By inverting eq.(56) with F substituting E one can also define an effective (relative) input/output Hamming distance \mathcal{D} at temperature t :

$$\mathcal{D}(t, d) = \frac{2}{\pi} \arccos e^{\frac{-F(t, d)}{2}}.\tag{61}$$

This corresponds exactly to representing the recognition probability of the average memory as

$$(P_b^{\text{rec}})_{\text{av}} = \cos^{2b} \left(\frac{\pi}{2} \mathcal{D}(b, d) \right),\tag{62}$$

which can also be taken as the primary definition of the effective Hamming distance.

The function $\mathcal{D}(b, d)$ provides a complete description of the behaviour of the average probabilistic quantum associative memory with a minimal distance Hamming distance d .

This can be used to tune its performance. Indeed, suppose that one wants the memory to recognize and identify inputs with up to ϵn corrupted inputs with an efficiency of ν ($0 \leq \nu \leq 1$). Then one must choose a number b of control qbits sufficiently large that $(\mathcal{D}(b, \epsilon n) - \epsilon) \leq (1 - \nu)$ and a threshold T of repetitions satisfying $T \geq 1/\cos^{2b}(\frac{\pi}{2}\mathcal{D}(b, \epsilon n))$, as illustrated in Fig. 2 below.

A first hint about the general behaviour of the effective distance function $\mathcal{D}(b, d)$ can be obtained by examining closer the energy eigenvalues (56). For small Hamming distance to the input these reduce to

$$E^k \simeq \frac{\pi^2}{4} \left(\frac{d_H(i, p^k)}{n} \right)^2, \quad \frac{d_H(i, p^k)}{n} \ll 1. \quad (63)$$

Choosing again the normalization in which $|i\rangle = |0 \dots 0\rangle$ and introducing a “spin” s_i^k with value $s_i^k = -1/2$ if qbit i in pattern p^k has value $|0\rangle$ and $s_i^k = +1/2$ if qbit i in pattern p^k has value $|1\rangle$, one can express the energy levels for $d_H/n \ll 1$ as

$$E^k = \frac{\pi^2}{16} + \frac{\pi^2}{4n^2} \sum_{i,j} s_i^k s_j^k + \frac{\pi^2}{4n} \sum_i s_i^k. \quad (64)$$

Apart from a constant, this is the Hamiltonian of an infinite-range antiferromagnetic Ising model in presence of a magnetic field. The antiferromagnetic term favours configurations k with half the spins up and half down, so that $s_{\text{tot}}^k = \sum_i s_i^k = 0$, giving $E^k = \pi^2/16$. The magnetic field, however, tends to align the spins so that $s_{\text{tot}}^k = -n/2$, giving $E^k = 0$. Since this is lower than $\pi^2/16$, the ground state configuration is ferromagnetic, with all qbits having value $|0\rangle$. At very low temperature (high b), where the energy term dominates the free energy, one expects thus an ordered phase of the quantum associative memory with $\mathcal{D}(t, d) = d/n$. This corresponds to a perfect identification of the presented input. As the temperature is raised (b decreased) however, the thermal energy embodied by the entropy term in the free energy begins to counteract the magnetic field. At very high temperatures (low b) the entropy approaches its maximal value $S(t = \infty) = 0$ (with the normalization chosen here). If this value is approached faster than $1/t$, the free energy will again be dominated by the internal energy. In this case, however, this is not any more determined

by the ground state but rather equally distributed on all possible states, giving

$$\begin{aligned} F(t = \infty) &= U(t = \infty) = \frac{-1}{1 - \frac{d}{n}} \int_{\frac{d}{n}}^1 dx \, 2 \log \cos \left(\frac{\pi}{2} x \right) \\ &= \left(1 + \frac{d}{n} \right) 2 \log 2 + O \left(\left(\frac{d}{n} \right)^2 \right), \end{aligned} \quad (65)$$

and leading to an effective distance

$$\mathcal{D}(t = \infty, d) = \frac{2}{3} - \frac{2 \log 2}{\pi \sqrt{3}} \frac{d}{n} + O \left(\left(\frac{d}{n} \right)^2 \right). \quad (66)$$

This value corresponds to a disordered phase with no correlation between input and output of the memory.

A numerical study of the thermodynamic potentials in (60) and (61) indeed confirms a phase transition from the ordered to the disordered phase as the effective temperature is raised. In Fig. 2 we show the effective distance \mathcal{D} and the entropy S for 1 Mb ($n = 8 \times 10^6$) patterns and $d/n = 1\%$ as a function of the inverse temperature b (the entropy is rescaled to the interval $[0,1]$ for ease of presentation). At high temperature there is indeed a disordered phase with $S = S_{\max} = 0$ and $\mathcal{D} = 2/3$. At low temperatures, instead, one is in the ordered phase with $S = S_{\min}$ and $\mathcal{D} = d/n = 0.01$. The effective Hamming distance plays thus the role of the order parameter for this quantum phase transition.

The phase transition occurs around $b_{\text{cr}} \simeq 10^{-1}$. The physical regime of the quantum associative memory ($b = \text{positive integer}$) lies thus just above this transition. For a good accuracy of pattern recognition one should choose a fictitious temperature low enough to be well into the ordered phase. As is clear from Fig. 2, this can be achieved already with a number of control qubits $b = O(10^4)$.

Having described at length the information retrieval mechanism for complete, but possibly corrupted patterns, it is easy to incorporate also incomplete ones. To this end assume that only $q < n$ qubits of the input are known and let us denote these by the indices $\{k1, \dots, kq\}$. After assigning the remaining qubits randomly, there are two possibilities. One can just treat the resulting complete input as a noisy one and proceed as above or, better, one can limit the operator $(d_H)_m$ in the Hamiltonian (42) to

$$(d_H)_m = \sum_{i=1}^q \left(\frac{\sigma^z + 1}{2} \right)_{m_{ki}}, \quad (67)$$

Order-Disorder Transition of QAMs

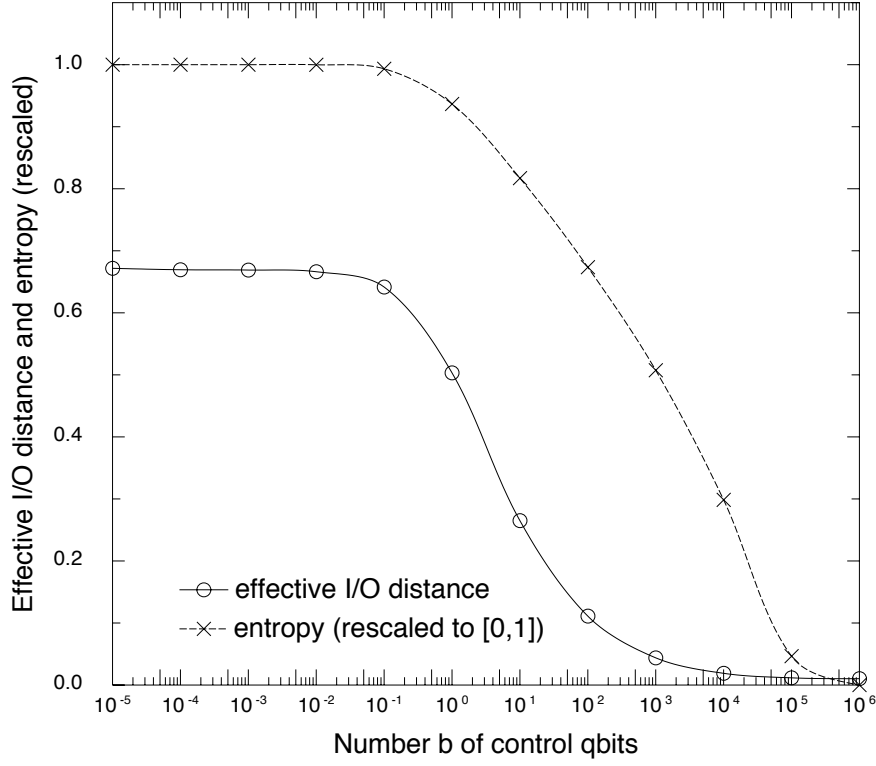


FIG. 2: Effective input/output distance and entropy (rescaled to $[0,1]$) for 1Mb patterns and $d/n = 1\%$.

so that the Hamming distances to the stored patterns are computed on the basis of the known qubits only. After this, the pattern recall process continues exactly as described above. This second possibility has the advantage that it does not introduce random noise in the similarity measure but it has the disadvantage that the operations of the memory have to be adjusted to the inputs.

Finally, it is fair to mention that the model of probabilistic quantum associative memory presented here has been criticised [36] on three accounts:

- It has been claimed that the same result could have been obtained by storing only one of the p patterns in n classical bits and always using this single pattern as the same

output independently of the input, provided the input has a Hamming distance to the unique stored pattern lower than a given threshold, otherwise the input would not be recognized.

- It has been claimed that the Holevo theorem bounds the number of patterns that can be stored in a quantum associative memory.
- It has been pointed out that the complexity of memory preparation prevents the efficient storing of patterns.

This criticism is wrong on the first two accounts and partially justified on the third [37]. It is true that both the quantum memory and the proposed equivalent classical prescription are based on probabilistic recognition and identification processes. In the proposed classical alternative, however the probabilities for both recognition and identification depend on *one unique, fixed and random pattern* whereas in the quantum memory, exactly due to its quantum character, these probabilities depend on *the distribution of all stored patterns*. These probabilities are such that an input different from most stored patterns is more difficult to recognize than an input similar to many stored memories and that the identification probability distribution can be peaked with any prescribed accuracy on the stored pattern most similar to the input. In the proposed classical alternative, given that only one single pattern can be stored on the n classical bits, the recognition or lack thereof depend on the distance to a randomly chosen pattern and the identification probability is a delta function peaked on this fixed random pattern. In other words there is no correlation whatsoever between input and output apart from the fact that they have Hamming distance below a certain threshold, a prescription that can hardly qualify as an associative memory: it would indeed be a boring world the one in which every stimulus would produce exactly the same response, if any response at all. Also, the Holevo theorem [26] does not impose any limitation on this type of probabilistic quantum memories. The Holevo theorem applies to the situation in which Alice codes information about a classical random variable in a quantum state and Bob tries to retrieve the value of this random variable by measurements on the received quantum state. In the present case Alice gives to Bob also corrupted or incomplete *classical information* about the random variable (the input) and Bob can use also a *unitary transformation that encodes both the memories and the input* (operator $R(i)$ in (47)) in addition to measurements, a completely different situation. Contrary to what the

authors of [36] affirm, a memory that "knows the patterns it is supposed to retrieve" not only makes sense but it is actually the very definition of an associative memory: if the memory would not "know" the data it has to retrieve it would just be a random access database, exactly the architecture that one wants to improve by content association, the mechanism whose goal is to recognize and correct corrupted or incomplete inputs. The dynamics of the classical Hopfield model "knows" the patterns it is supposed to retrieve: they are encoded in the neuronal weights. So does any human brain. Finally, the third critique is partially correct. The complexity of the memory operator M is $O(pn)$ and thus the original claim [13] of an exponential capacity gain by quantum associative memories is excessive. This, however, does not invalidate the main claim, a large gain in capacity is made possible by quantum mechanics, albeit only a polynomial one. This correction has been incorporated in the present review.

C. Efficiency, complexity and memory tuning

In this last section we would like to address the efficient implementation of probabilistic quantum memories in the quantum circuit model [4] and their accuracy tuning.

We have stressed several times that all unitary operators involved in the memory preparation can be realized as a sequence of one- and two-qubit operators. It remains to prove that this is true also for pattern retrieval and that all these operators can be implemented in terms of a small set of universal gates. To this end we would like to point out that, in addition to the standard NOT, H (Hadamard), XOR, 2XOR (Toffoli) and nXOR gates [4] we have introduced only the two-qubit gates CS^i in eq. (22) and the unitary operator $\exp(i\pi\mathcal{H}/2n)$. The latter can, however also be realized by simple gates involving only one or two qubits. To this end we introduce the single-qubit gate

$$U = \begin{pmatrix} e^{i\frac{\pi}{2n}} & 0 \\ 0 & 1 \end{pmatrix}, \quad (68)$$

and the two-qubit controlled gate

$$CU^{-2} = |0\rangle\langle 0| \otimes 1 + |1\rangle\langle 1| \otimes U^{-2}. \quad (69)$$

It is then easy to check that $\exp(i\pi\mathcal{H}/2n)$ in eq. (41) can be realized as follows:

$$e^{i\frac{\pi}{2n}\mathcal{H}} |\psi_2\rangle = \prod_{i=1}^n (CU^{-2})_{cm_i} \prod_{j=1}^n U_{m_j} |\psi_2\rangle , \quad (70)$$

where c is the control qbit for which one is currently repeating the algorithm. Essentially, this means that one implements first $\exp(i\pi d_H/2n)$ and then one corrects by implementing $\exp(-i\pi d_H/n)$ on that part of the quantum state for which the control qbit $|c\rangle$ is in state $|1\rangle$.

Using this representation for the Hamming distance operator one can count the total number of simple gates that one must apply in order to implement one step of the information retrieval algorithm. This is given by $(6n+2)$ using the auxiliary register for the input and by $(4n+2)$ otherwise. This retrieval step has then to be repeated for each of the b control qbits. Therefore, implementing the projection by repeated measurements, the overall complexity C of information retrieval is bounded by

$$C \leq Tb(6n+2)C_M , \quad (71)$$

where C_M is the complexity of the memory preparation, given by the operator M or a probabilistic cloning machine. In particular, it is given by

$$C = Tb(6n+2)(p(2n+3)+1) , \quad (72)$$

for the simplest version of the algorithm, using memory preparation by M and an auxiliary input register.

The computation of the overall complexity is easier for the information retrieval algorithm which uses the amplitude amplification technique. In this case the initial memory is prepared only once by a product of the operators M , with complexity $p(2n+3)+1$ and $R(i)$, with complexity $b(4n+2)$. Then one applies T times the operator Q , with complexity $p(4n+6)+b(8n+4)+2+C_S+C_{S_0}$, where C_S and C_{S_0} are the polynomial complexities of the oracles implementing S and S_0 . This gives

$$\begin{aligned} C &= T[p(4n+6)+b(8n+4)+2+C_S+C_{S_0}] + \\ &+ p(2n+3)+b(4n+2)+1 . \end{aligned} \quad (73)$$

As expected, the memory complexity (be it (72) or (73)) depends on both T and b , the parameters governing the recognition and identification efficiencies. The major limitation

comes from the factor p representing the total number of stored patterns. Note however that, contrary to classical associative memories, one can efficiently store and retrieve any polynomial number of patterns due to the absence of spurious memories and crosstalk.

Let us finally show how one can tune the accuracy of the quantum memory. Suppose one would like to recognize on average inputs with up to 1% of corrupted or missing bits and identify them with high accuracy. The effective i/o Hamming distance \mathcal{D} shown in Fig. 2 can then be used to determine the values of the required parameters T and b needed to reach this accuracy for the average memory. For $b = 10^4$ e.g., one has $\mathcal{D} = 0.018$, which gives the average i/o distance (in percent of total qbits) if the minimum possible i/o distance is 0.01. For this value of b the recognition probability is $3.4 \cdot 10^{-4}$. With the measurement repetition technique one should thus set the threshold $T \simeq 3000$. Using amplitude amplification, however, one needs only around $T = 54$ repetitions. Note that the values of b and T obtained by tuning the memory with the effective i/o Hamming distance become n -independent for large values of n . This is because they are intensive variables unaffected by this "thermodynamic limit". For any fixed p polynomial in n , the information retrieval can then be implemented efficiently and the overall complexity is determined by the accuracy requirements via the n -independent parameters T and b .

V. CONCLUSION

We would like to conclude this review by highlighting the fundamental reason why a probabilistic quantum associative memory works better than its classical counterpart and pointing out about some very intuitive features of the information retrieval process.

In classical associative memories, the information about the patterns to recall is typically stored in an energy function. When retrieving information, the input configuration evolves to the corresponding output, driven by the dynamics associated with the memory function. The capacity shortage is due to a phase transition in the statistical ensemble governed by the memory energy function. Spurious memories, i.e. spurious metastable minima not associated with any of the original patterns become important for loading factors p/n above a critical value and wash out completely the memory, a phenomenon that goes by the name of crosstalk. So, in the low p/n phase the memory works perfectly in the sense that it outputs always the stored pattern which is most similar to the input. For p/n above the

critical value, instead, there is an abrupt transition to total amnesia caused by spurious memories.

Probabilistic quantum associative memories work better than classical ones since they are *free from spurious memories*. The easiest way to see this is in the formulation

$$|m\rangle = M |0\rangle . \quad (74)$$

All the information about the stored patterns is encoded in the unitary operator M . This generates a quantum state in which all components that do not correspond to stored patterns have exactly vanishing amplitudes.

An analogy with the classical Hopfield model [5] can be established as follows. Instead of generating the memory state $|m\rangle$ from the initial zero state $|0\rangle$, one can start from a uniform superposition of the computational basis. This is achieved by the operator MW defined by

$$\begin{aligned} |m\rangle &= MW \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} |j\rangle , \\ W &\equiv \prod_{j=1}^n H_j . \end{aligned} \quad (75)$$

Now, this same result can also be obtained by Grover's algorithm, or better by its generalization with zero failure rate [38]. Here the state $|m\rangle$ is obtained by applying to the uniform superposition of the computational basis q times the search operator X defined in

$$\begin{aligned} |m\rangle &= X^q \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} |j\rangle , \\ X &\equiv -W J_0 W J , \end{aligned} \quad (76)$$

where J rotates the amplitudes of the states corresponding to the patterns to be stored by a phase ϕ which is very close to π (the original Grover value) for large n and J_0 does the same on the zero state. Via the two equations (75) and (76), the memory operator M provides an implicit realization of the phase shift operator J . Being a unitary operator, this can always be written as an exponential of an hermitian Hamiltonian \mathcal{H}_M , which is the quantum generalization of a classical energy function. By defining $J \equiv \exp(-i\mathcal{H}_M)$ one obtains an energy operator which is diagonal in the computational basis and such that the patterns to be stored have energy eigenvalues $E = -\phi \simeq -\pi$ while all others have energy eigenvalues $E = 0$. This formulation is the exact quantum generalization of the Hopfield

model; the important point is that the operator M realizes efficiently a dynamics in which the patterns to be stored are always, for any number p of patterns, the exact global minima of a quantum energy landscape, without the appearance of any spurious memories.

The price to pay is the probabilistic nature of the information retrieval mechanism. As always in quantum mechanics, the dynamics determines only the evolution of probability distributions and the probabilistic aspect is brought in by the collapse of this probability distributions upon measurement. Therefore, contrary to the classical Hopfield model in the low p/n phase, one does not always have the absolute guarantee that an input is recognized and identified correctly as the stored pattern most similar to the input, even if this state has the highest probability of being measured. But, after all, this is a familiar feature of the most concrete example of associative memory, our own brain, and should thus not be so disturbing. Indeed, it is not only the probabilistic nature of information retrieval that is reminiscent of the behaviour of the human brain but also the properties of the involved probability distributions. These are such that inputs very similar to a cluster of stored patterns will be much easier to recognize than inputs farther away from all stored memories, although the former situation will lead to a more difficult identification of the most similar memory. Quantum entanglement allows to construct such a probabilistic content association with a very high storage capacity.

-
- [1] See e.g.: Jeff Hawkins (with S. Blakeslee), *On Intelligence*, Times Books (2004), R. Kurzweil, *How To Create A Mind*, Penguins Books, London (2012) and references therein.
 - [2] See e.g.: C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer Verlag, Singapore (2006).
 - [3] See e.g.: M. Davis, *Engines of Logic: Mathematicians and the Origin of the Computer*, W. Norton Company, New York (2000) and references therein.
 - [4] For a review see: M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge (2000); A. O. Pittenger, *An Introduction to Quantum Computing Algorithms*, Birkhäuser, Boston (2000).
 - [5] For a review see: B. Müller and J. Reinhardt, *Neural Networks*, Springer-Verlag, Berlin (1990); T. Kohonen, *Self-Organization and Associative Memory*, Springer-Verlag, Berlin (1984).

- [6] For a review see: L. R. Rabiner, *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*, Proc. of the *IEEE* **77** (1989) 257.
- [7] J.J. Hopfield, Proc. Natl. Acad. Scie. USA **79**, 2554 (1982).
- [8] B. Kosko, IEEE Trans. on Systems, Man and Cybernetics **18**, 49 (1988).
- [9] See e.g.: H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press, Oxford (2001).
- [10] See e.g. M. Mezard, G. Parisi and M. A. Virasoro, *Spin Glass Theory and Beyond*, World Scientific, Singapore (1987).
- [11] P. W. Shor, SIAM J. Computing **26** (1997) 1484.
- [12] L. Grover, Phys. Rev. Lett. **79** (1997) 325.
- [13] C. A. Trugenberger, *Phys. Rev. Lett.* **87** (2001) 067901, see also P. Ball, "Brain Inspires New Memories" *Nature News*, August 6 (2001).
- [14] C. A. Trugenberger, *Phys. Rev. Lett.* **89** 2002 277903.
- [15] For a review see: C. A. Trugenberger, "Quantum Pattern Recognition", *Quantum Inf. Process.* **1** (2002) 471.
- [16] M. Cristina Diamantini and C. A. Trugenberger, *Phys. Rev. Lett.* **97** (2006) 130503.
- [17] M. Sasaki, A. Carlini and R. Jozsa, Phys. Rev. **A64** (2001) 022317; M Sasaki and A. Carlini, *Phys. Rev.* **A66** (2002) 022303 ; R. Schützhold, *Phys. Rev.* **A67** (2003) 062311.
- [18] N. Surlas, Nature **339**, 693 (1989); I. Kanter and D. Saad, Phys. Rev. Lett. **83**, 2660 (1999); Y. Kabashima, T. Murayama and D. Saad, Phys. Rev. Lett. **84**, 1355 (2000).
- [19] Y. Kabashima, T. Murayama and D. Saad, Phys. Rev. Lett. **84** 2030 (2000).
- [20] W. S. McCulloch and W. Pitts, *Bull. Math. Biophys.* **5** (1943) 115
- [21] O. Mandel, M. Greiner, A. Widera, T. Rom, T. W. Hänsch and I. Bloch, *Nature* **425** (2003) 937.
- [22] B. E. Kane, *Nature* **393** (2003) 133.
- [23] See e.g. G. Parisi, *Statistical Field Theory*, Addison-Wesley, Redwood City (1988).
- [24] See e.g.: S. Sachdev, *Quantum Phase Transitions*, Cambridge University Press, Cambridge (1999).
- [25] See e.g. A. Barenco, C. Bennet, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin and H. Weinfurter, *Phys. Rev.* **A52**, 3457 (1995).
- [26] A. S. Holevo, *Problems of Information Transmission* **9** (1973) 177.

- [27] D. Ventura and T. Martinez, *Found. Phys. Lett.* **12** (1999) 547.
- [28] Y. Tanaka, T. Ichikawa, M. Tada-Umezaki, Y. Ota and M. Nakahara, *Int. J. Quant. Info.* **9** (2011) 1363.
- [29] W. Wootters and W. Zurek, *Nature* 299, 802 (1982).
- [30] V. Buzek and M. Hillery, *Phys. Rev.* **A54**, 1844 (1996).
- [31] N. Gisin and S. Massar, *Phys. Rev. Lett.* **79**, 2153 (1997); D. Bruss, A. K. Ekert and C. Macchiavello, *Phys. Rev. Lett.* **81**, 2598 (1998).
- [32] D. Bruss, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello and J. A. Smolin, *Phys. Rev.* **57** (1998) 2368.
- [33] L.-M. Duan and G.-C. Guo, *Phys. Rev. Lett.* **80**, (1998) 4999.
- [34] A. Chefles and S. M. Barnett, *Phys. Rev.* **A60**, **136** (1999) 136.
- [35] G. Brassard, P. Hoyer, M. Mosca and A. Tapp, *Amplitude Amplification and Estimation*, quant-ph/0005055.
- [36] T. Brun, H. Klauck, A. Nayak, M. Roetteler and Ch. Zalka, *Phys. Rev. Lett.* **91** (2003) 209801.
- [37] C. A. Trugenberger, *Phys. Rev. Lett.* **91** (2003) 209802.
- [38] G. L. Long, *Phys. Rev.* **A64** (2001) 022307.